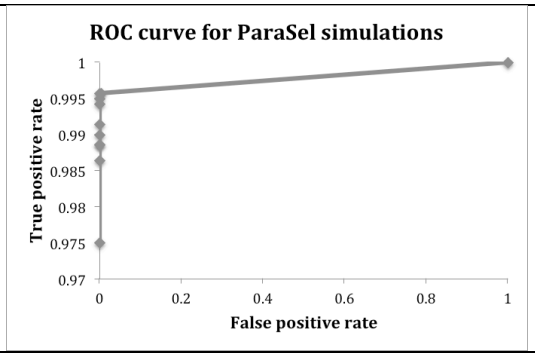
Table S2. Related to Methods. Performance results of the ParaSel model on simulated data. All datasets were simulated based on the relevant cVDPV2 phylogeny and branch lengths. Parameters of the HKY model were taken from the results of the cVDPV2 analysis. 100 datasets were simulated in each category; due to computational load, number of loci was reduced to 200. A receiving operator characteristic (ROC) curve is shown for different posterior probability thresholds (0 through 1, in steps of 0.1), for scenario 3 simulating parallel selection. The results show a false positive rate of zero for all thresholds of the posterior probability except when it equals zero (upper right corner). For the empirical analysis we remained with a highly conservative posterior threshold of 0.8, but report sites under a lower threshold in table S3.

#	Scenario simulated	Simulated data	Proportion of datasets identified under parallel selection[1]
1	Neutral nucleotide evolution	Nucleotide alignments simulated under the neutral HKY model	0%
2	Neutral nucleotide evolution, with an epistatic interaction between two sites	Nucleotide alignments simulated under the neutral HKY model. An epistatic interaction was simulated by accepting only alignments that allowed G:C, A:U, or G:U base-pairing at a given pair of sites	0%
3	Parallel selection operating on 10% of the alignment	Nucleotide alignments simulated based on the empirical cVDPV2 data. 10% of sites were simulated under the ParSel model with selection for "A", and 90% of the sites were simulated under the HKY model.	100%
		Percent sites identified as parallel selection posterior probability >= 0.8	69.6%
		Percent sites identified as parallel selection, non-ancestral "A" [2]	98.8%
			

[1] (ΔAIC_c > 6, corresponding to P < 0.05)

[2] Sites with ancestral "A" state remain "A" throughout the entire simulation under selection and hence are detected under the neutral model as slow evolving sites.

Table S3. Related to Fig. 3. (upper panel) AIC values and parameters for the analysis of the different genomic nucleotide regions under the ParaSel model and null model. (lower panel) Substitutions reported by ParaSel with intermediate posterior probabilities (0.1-0.8). Substitutions with a posterior probability higher than 0.8 are listed in Fig. 3.

Genomic region	**ParaSel model AIC**	**Null model AIC**	**P-value**	**Maximum likelihood parameter estimates (Methods)**
Capsid	82,475.2	82,817.6	$4x10^{-75}$	P=0.007 S=10 α=0.29 κ=10.8 β=0.0001 τ=0.97
5' UTR	2,634	2,831.2	$1.5x10^{-43}$	P=0.02 S=8.9 α=1.02 κ=10.2 β=8e-06 τ=1

Genomic position number	**Substitution**	**Posterior probability for parallel selection**	**Type of substitution**	**Disrupts CpG/Upa?**
1353	T→A	0.2	Synonymous	-
1443	T→C	0.12	Synonymous	UpA
1605	A→T	0.18	Synonymous	UpA
1869	A→C	0.13	Synonymous	UpA
1997	A→G	0.46	Non-syn.	-
1999	T→A	0.29	Non-syn.	-
2280	C→T	0.15	Synonymous	-
2443	C→A	0.12	Non-syn.	CpG
2697	A→T	0.14	Synonymous	-
2779	C→A	0.29	Non-syn.	CpG
2931	A→T	0.54	Synonymous	-
3000	T→C	0.23	Synonymous	UpA
3147	A→T	0.13	Synonymous	-
3210	C→A	0.58	Synonymous	CpG
3294	A→T	0.14	Synonymous	-

Table S4. **Related to Fig. 6. Sites present in HEV-C recombination partner that differ from OPV2**. Only sites that are present in more than 90% of the HEV-C partners and differ from OPV2 are shown. Changes which may lead to a functional change are marked: two non-conservative amino-acid changes, two sites that form part of known secondary structures (Burrill et al., 2013), and sites that lead to a disruption of a CpG or UpA dinucleotide.

OPV2 locus	OPV2 codon	HEV-C codon	OPV2 AA	HEV-C AA	Mutation type	Part of secondary structure	Disrupts CpG/UpA
4546	tta	ttg	L	L	SYN	-	+
4717	gac	gat	D	D	SYN	-	+
4828	gcc	gca	A	A	SYN	-	-
4879	gcg	gcc	A	A	SYN	-	+
4918	gtc	gta	V	V	SYN	-	-
4930	tac	tat	Y	Y	SYN	-	-
5011	tgt	tgc	C	C	SYN	-	-
5068	aga	agg	R	R	SYN	-	-
5155	cag	caa	Q	Q	SYN	-	-
5221	aac	aat	N	N	SYN	-	+
5227	ttg	ctg	L	L	SYN	-	-
5254	gtg	gta	V	V	SYN	-	-
5377	ggt	gga	G	G	SYN	-	-
5380	gtc	gtg	V	V	SYN	-	+
5407	gct	gca	A	A	SYN	-	-
5428	act	aca	T	T	SYN	-	-
5443	aaa	aag	K	K	SYN	-	-
5659	gcc	gct	A	A	SYN	-	-
5662	aaa	aag	K	K	SYN	-	-
5689	aat	aac	N	N	SYN	-	-
5827	ccc	cct	P	P	SYN	+[3]	+
5956	att	atc	I	I	SYN	+[3]	-
6028	att	agt	I	S	NON-SYN[1]	-	-
6178	agg	aga	R	R	SYN	-	-
6187	aca	gtg	T	V	NON-SYN[2]	-	-
6271	cat	cac	H	H	SYN	-	-
6286	ctc	tta	L	L	SYN	-	-
6358	gca	gcc	A	A	SYN	-	-
6364	gac	gat	D	D	SYN	-	-
6409	aaa	aag	K	K	SYN	-	-
6430	aag	aaa	K	K	SYN	-	-
6448	aag	aaa	K	K	SYN	-	-
6466	tta	ctg	L	L	SYN	-	+
6625	ttt	ttc	F	F	SYN	-	-
6670	gat	gac	D	D	SYN	-	-
6736	aca	acc	T	T	SYN	-	-
6754	ctc	ttg	L	L	SYN	-	-
6874	tgc	tgt	C	C	SYN	-	+
6880	aaa	aag	K	K	SYN	-	-
6937	aac	aat	N	N	SYN	+	-
7294	gcc	gct	A	A	SYN	-	-
7303	aac	aat	N	N	SYN	-	+
7354	att	atc	I	I	SYN	-	-

[1] Nonpolar AA -> Polar, uncharged AA (3D protein)
[2] Polar, uncharged AA -> Nonpolar AA (3D protein)
[3] Part of the RNase L ciRNA

Table S5. **Related to Fig. 4.** Sequence coverage and number of variant alleles detected at each passage using CirSeq approach.

Passage	Average number of reads (coverage) per locus	Sites where a minor allele was detected	Number of variant alleles	Total number of point mutations
P1	127,951	7,266	17,122	190,662
P2	259,374	6,088	18,755	568,215
P3	342,821	7,383	20,230	624,655
P4	172,317	7,357	17,303	507,590
P5	205,665	7,317	18,682	379,753
P6	312,154	7,385	20,007	789,769
P7	272,699	7,391	19,880	666,253
Avg.	**241,854**	**7,170**	**18,854**	**532,414**